eposcorr

February 23, 2011

Abstract

Adds corrected source positions to the EPIC source list by cross-correlation with optical source catalogue.

1 Instruments/Modes

Instrument	Mode
EPIC	Imaging

2 Use

pipeline processing	yes
interactive analysis	no

3 Description

The **eposcorr** task reads the X-ray source positions from an input fits-file and correlates those with the positions from an optical source catalogue. The correlation algorithm checks whether there are offsets in RA and DEC which optimize the correlation. To do this a particular statistic is used and its value calculated for a grid of 30x30 offset values with a resolution, which can be controled by input parameters. If demanded a similar grid search is done for roll angle errors. In that case, first a 3dimensional grid search will be done and then several iterations with alternating RA/DEC offsets and rotational offsets are made. The optimum offsets are then used to correct the input source positions which will be added as seperate columns to input X-ray source catalogue.

3.1 Statistical method

The statistic for optimizing the match between optical and X-ray sources is:

$$L = \sum_{i=1}^{n_x} \sum_{j=1}^{n_o} \exp(-\frac{1}{2}(\frac{r_{ij}}{\sigma_{ij}})^2), \quad (1)$$

with r_{ij} the distance between an X-ray (i) and an optical source (j), σ_{ij} the associated error and n_x resp. n_o the number of X-ray sources and optical sources in the list. In the eposcorr task only those optical sources are considered which are within 5σ of an X-ray source (for a given position offset).

In the following we will assume that the errors in the RA and DEC are equal and uncorrelated and follow a gaussian distribution, we can then write $\sigma_x = \sigma_y = \sigma$[1]. For the expectation value of L for a given X-ray source we get:

$$L = \frac{1}{2\pi\sigma^2}\int dx \int dy \exp(-\frac{1}{2}\frac{x^2+y^2}{\sigma^2})\ \exp(-\frac{1}{2}\frac{x^2+y^2}{\sigma^2}) = \frac{1}{2\pi\sigma^2}\int dr\ 2\pi r\ \exp(-\frac{r^2}{\sigma^2}) = \frac{1}{2}. \quad (2)$$

The associated variance in L is:

$$< L^2 > - < L >^2 = \int dr\ 2\pi r\ \exp(-\frac{1}{2}\frac{r^2}{\sigma^2})\Big(\exp(-\frac{1}{2}\frac{r^2}{\sigma^2})\Big)^2 - \Big(\frac{1}{2}\Big)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \quad (3)$$

Of course, in practice there will be chance coincidences. For chance coincidences the chance that an optical counter part will be within a distance r from the source is $\pi r^2/\pi(5\sigma)^2$ (i.e. within an error circle of 5σ). This gives for the expected value L for a chance coincidence:

$$L = \int_0^{5\sigma} dr \frac{2\pi r}{\pi 25\sigma^2}\exp(-\frac{1}{2}\frac{r^2}{\sigma^2}) = \frac{2}{25}\Big(1-\exp(-\frac{25}{2})\Big) \simeq \frac{2}{25}. \quad (4)$$

And for the variance in L:

$$< L^2 > - < L >^2 = \int_0^{5\sigma} dr \frac{2\pi r}{\pi 25\sigma^2}\exp(-\frac{r^2}{\sigma^2}) - \Big(\frac{2}{25}\Big)^2 = \frac{1}{25}\Big(1-\exp(-25)\Big) - \Big(\frac{2}{25}\Big)^2 \simeq \frac{1}{25} - \Big(\frac{2}{25}\Big)^2 = \frac{21}{25^2}. \quad (5)$$

The number of chance coincidences can be estimated using poisson statistics with a poisson parameter of $\mu = 25\pi\sigma^2\lambda$, with λ the average number of optical sources per unit area. The expected number of sources is thus $\sum_{i=1}^{n_x}\mu_i$ (where the subsript i denotes values for each X-ray source, thus allowing for fluctuations in the number of optical sources per area). The expected value for L is:

$$L = \sum_{i=1}^{n_x}\mu_i \int_0^{5\sigma_i} dr\ \frac{2\pi r}{\pi\ 25\sigma^2}\exp(-\frac{1}{2}\frac{r^2}{\sigma^2}) = \frac{2}{25}\sum_{i=1}^{n_x}\mu_i. \quad (6)$$

How many counterparts do we need to discriminate between chance coincidences or real counter parts? This question is not easy to answer, as **eposcorr** optimizes L and also for the number of counter parts. This means that poissonian statistics may bnot be valid. To get at least an approximate answer, we equate:

$$(L_{exp} - 2\sigma_L)_{gaussian} = (L_{exp} + 2\sigma_L)_{poissonian}, \quad (7)$$

or,

$$\frac{N}{2} - 2\sqrt{\frac{N}{12}} = \frac{2N}{25} + 2\sqrt{\frac{2N}{25^2}}. \quad (8)$$

The solution of this equation is $N = 5.1$. I therefore propose to use this number plus the number of degrees of freedom as the minimum threshold for accepting a result of **eposcorr**. This means that when offsets and in RA and DEC are corrected for the minimum number of optical counter parts should be 7, including a rotational correction this will be 8. This number will be contained in the keyword `NMATCHES`.

[1] This is different from the definition used by the source detect codes there $\sigma = \sqrt{\sigma_x^2 + \sigma_y^2}$. This is corrected for in **eposcorr**

3.2 The errors in the measured off-sets

At the moment we have no rigurous mathematical method to handle the statistical errors of the offsets found. For the moment we use the square root of the second order moment in L as a function of off-set divided by the number of optical counter parts.

3.3 Some spherical geometry

eposcorr calculates angular distances directly using spherical trigonimetry. In particular the following is used to calculate the angular distance, θ, between two sources (labeled 1 and 2).

$$\cos(\theta) = \sin(\delta_1)\sin(\delta_2) + \cos(\delta_1)\cos(\delta_2)cos(\Delta\alpha), \tag{9}$$

where δ refers to the declination of a source and $\Delta\alpha = \alpha_1 - \alpha_2$, the difference of the right ascension of the two sources.

4 Parameters

This section documents the parameters recognized by this task (if any).

Parameter	Mand	Type	Default	Constraints
xrayset	yes	set		

File name for the fits binary table containing X-ray source positions and their errors.

opticalset	yes	set		

File name for the fits binary table containing optical source positions and their errors.

findrotation	no	boolean	no	

Search for a rotational offset or not?

niter	no	integer	3	1–8

Once a solution is found a new iteration can be made using a smaller grid size. This parameter determines the number of iterations.

maxoffset	no	real	8.0	4.0–15.0

Constraint on the maximum RA/DEC offset.

maxrotation	no	real	1.0	0.5–4.0

Constraint on the maximum rotational offset.

maxdist	no	real	20.0	10.0–120.0

All optical sources with distances > `maxdist` will be removed from the source list, in order to speed up the correlation.

maxposnerr	no	real	20.0	1.0–120.0

neglect xray sources with positional errors > `maxposnerr` in position fitting in order to get a closer match to the real positions on the sky.

usemaxdist	no	boolean	yes	

Use `maxdist` to make a preselection of optical sources?

minbmagn	no	real	0.0	-2.0–30.0

Preselect the optical list on basis of B Magnitude. Here the mininum is set.

maxbmagn	no	real	26.0	-2.0–30.0

Preselect the optical list on basis of B Magnitude. Here the maximum is set.

usebmagn	no	boolean	no	

Use the B magnitude to make a preselection of optical sources?

bmag	no	string	Bmagn	

name of B magnitude column

minrmagn	no	real	0.0	-2.0–30.0

Preselect the optical list on basis of R Magnitude. Here the mininum is set.

maxrmagn	no	real	26.0	-2.0–30.0

Preselect the optical list on basis of R Magnitude. Here the maximum is set.

usermagn	no	boolean	no	

Use the R magnitude to make a preselection of optical sources?

rmag	no	string	Rmagn	

name of R magnitude column

makeimage	no	boolean	no	

yes: the grid containing likelihood values will be written as a fits image file

imagename	no	filename	likelihood_test.fits	constraints

the name of the image file containing the likelihood values

opticalra	no	string	CAT_RA	

column name of optical catalogue position RA

opticaldec	no	string	CAT_DEC	

column name of optical catalogue position DEC

opticalradecerr	no	string	CAT_RADEC_ERR	

column name of optical catalogue position error

opticaltableext	no	string	RAWRES	

optical catalogue fits table extension

xrayra	no	string	RA	

column name of X-ray source position RA

xraydec	no	string	DEC	

column name of X-ray source position DEC

xrayradecerr	no	string	RADEC_ERR	

column name of X-ray source position error

xraytableext	no	string	SRCLIST	

X-ray source list fits table extension

calculateoffsets	no	boolean	yes	

calculate offsets; if no: use offsets/rotation provided by input parameter

raoffset	no	real	0.0	

RA offset (arcsec) to be used if caclulateoffsets=no

decoffset	no	real	0.0	

DEC offset (arcsec) to be used if calculateoffsets=no

rotation	no	real	0.0	

rotation angle (degrees) to be used if calculateoffsets=no

withmatchtable	no	boolean	no	

write optional table of matched objects

matchset	no	string	matches.fits	

name of matched objects file

rawxsyserr	no	real	1.5	

systematic error of input X-ray positions

minxsyserr	no	real	0.2	

irreducible systematic error of X-ray positions

maxsig	no	real	5.0	

sigma cutoff for X-ray/optical matches

usemaxsig	no	boolean	yes	

use sigma cutoff limit

5 Errors

This section documents warnings and errors generated by this task (if any). Note that warnings and errors can also be generated in the SAS infrastructure libraries, in which case they would not be documented here. Refer to the index of all errors and warnings available in the HTML version of the SAS documentation.

emptyList *(error)*
One of the source lists does not contain any data. (Can occurs during processing of the data.)

emptySrcList *(error)*
The data set with the X-ray source list is empty. (Input error.)

emptyCatList *(error)*
The data set with the optical source list is empty. (Input error.)

invalidDec *(warning)*
Apparently this source is very close to the equatorial pole and the RA is not well defined
corrective action: The value of `RA_CORR` for a particular source will be set to 0.

noBmagnColumn *(warning)*
The optical source list set does not contain a column with B magnitudes
corrective action: Ignore the B magnitudes, in case `usebmagn` was set to yes

noRmagnColumn *(warning)*
The optical source list set does not contain a column with R magnitudes
corrective action: Ignore the R magnitudes, in case `usermagn` was set to yes

noRaDcPnt *(warning)*
The X-ray data set does not contain the `RA_PNT` and/or `DEC_PNT` keywords. As its values are used for the rotation, without these kewyords no rotations can be found.
corrective action: Do not optimize for a rotatitional offset

invalidLikelihood *(warning)*
The value for the statistic used to optimize the offsets is too low, so that the solution may not be reliable.
corrective action: The corrective action is up to the user

NoUnit *(warning)*
No RA unit; deg assumed
corrective action: Assume RA values are in degrees

NoUnit *(warning)*
No DEC unit; deg assumed
corrective action: Assume DEC values are in degrees

NoUnit *(warning)*
No error unit; arcsec assumed
corrective action: Assume errors are in arcsec

6 Input Files

1. PPS product: list with right ascension (RA), declination (DEC). and position error (RA_DEC_ERR) of detected X-ray sources (default values of the column names are given in brackets)
2. Optical object catalogue used for cross-correlation, containing columns with right ascension (CAT_RA), declination (CAT_DEC), and position error (CAT_RADEC_ERR) of the catalogue objects (default values of the column names are given in brackets)

7 Output Files

From version 1.1 on the output routine tests if the columns RA_CORR and DEC_CORR are already present in the input/output file. Depending on the presence of these columns it will either add them or modify them. After the task has been run the following items will be present:

1. X-ray source list with two additional columns, being the corrected RA and DEC called RA_CORR and DEC_CORR.
2. The following keywords are added to the table: `RAOFFSET`, `DEOFFSET`, which indicate the offsets in RA and DEC used to correct the psoitions; `RAOFFERR`, `DEOFFERR` (the associated statistical errors) and `LIK_HOOD` (containing the likelihood statistic) and `NMATCHES`. The keywords `ROT_CORR` (rotional correction) and `ROT_ERR` (error in rotation) are also added. The keyword `LIK_NULL` contains the likelihood that can be expected from spurious matches only. It depends on the number of X-ray sources, the distribution of their positional errors, and the surface density of optical sources.
3. An optional image will be produced containing a map of likelihood values. This image can be used to check the uniqueness of the optimization routine.

8 Algorithm

```
subroutine eposcorr

    read x-ray and optical list

    call find_offset

    call correct_positions

    IF findrotation == TRUE THEN
      call find_all ! make a three dimensional search

    DO i=1, niter
      IF findrotation == TRUE THEN
         call find_rotation
      call find_offsets

     update the grid resolution
     ! for the last iteration a larger grid is used in order to get an error
     ! estimate
    END


    modify/write the x-ray list

end subroutine  eposcorr

subroutine find_offset

    x_shift= y_shift =0
    LOOP over x_offset grid elements
```

```
        LOOP over y_offset grid elements

           call subroutine likelihood
           store likelihood in array
           IF likelihood > best likelihood THEN
             update best likelihood, x_shift=x_offset , y_shift=y_offset

        END LOOP
    END LOOP

    xm = ym = 0
    total_likelihood = 0
    LOOP over likelihood array elements (= corresponding x_offset, y_offset)
      IF( distance( x/y_offset , x/y_shift ) < 0.5* grid size ) THEN
        xm = xm + likelihood * (x_offset - x_shift)
        ym = ym + likelihood * (y_offset - y_shift)
        total_likelihood = total_likelihood + likelihood
      END IF
    END LOOP

    ! New best offsets:
    x_shift = xm/total_likelihood
    y_shift = ym/total_likelihood

end subroutine find_offset

subroutine likelihood

    set likelihood = 0
    set nmatches   = 0
    LOOP over elements X-ray list
        LOOP over elements optical list

            calculate the combined position error (sigma)
            calculate the distance (d) between the optical and X-ray source
                      for given x and y offsets
            IF x-ray source within 5 sigma of optical source THEN
              nmatches =  nmatches + 1
              likelihood = likelihood + exp{ -0.5 * (d/sigma)^2 }
            END IF

        END LOOP
    END LOOP

end subroutine  likelihood
```

9 Comments

- At the moment the critical value for the likelihood is set to 0.3, which is somewhat arbitrary.

10 Future developments

- More research will be done on the expectation values for the likelihood statistic. This will include several possible situations (purely poisonian coincidences, real correlations and mixtures). This will help distinguishing reliable corrections from unreliable corrections. This implies that some additional keywords may be added in future which give information about the correlation.
- Some research could be done to other methods for determining offset. One can think of Fast Fourier Transforms for fast correlations, or a method based on triangle patterns.

References